DIPLOMARBEIT

Titel der Diplomarbeit

„The predictive power of leading economic indicators: An analysis in time and frequency domain"

Verfasser

Michael Wüger

angestrebter akademischer Grad

Magister der Sozial- und Wirtschaftswissenschaften (Mag.rer.soc.oec.)

Wien, im November 2010

Contents

List of Figures

List of Tables

1. Introduction

"German business confidence rises to three-year high in September:

German business confidence rose unexpectedly to a three-year high in September,

strengthening hopes that Europe's largest economy will not suffer a double-dip recession."

Telegraph (25th Sept. 2010)

Headlines like this can be found in press articles all over the world, whenever new values of leading economic indicators are released. In general, these leading indicators are measures which change before the economy is altered as a whole. Since Burns and Mitchell (1946) brought up the idea of economic indicator analysis in their paper "Measuring Business Cycles", the importance of leading indicators has grown rapidly throughout the past century. Nowadays the additional value of these measures is mostly accepted in the academic discussion and they are often used in empirical research. The analysis of leading economic indicators has become a main feature for identifying business cycles and the prediction of turning points within an economy's cycle. But leading indicators are not only important for the discipline of economic research. They are also well-known in public, because of the strong coverage in the media. The entire media landscape, from tabloid to scientific journal, is publishing the development of leading economic indicators and therefore these indicators play an important role for the expectation building process within the economy.

On the fundamentals of Burns and Mitchell (1946), a large number of institutes, organisations and universities have constructed different leading indicators, which are used to predict the upcoming activity in the economy. Some are designed to reflect the supply side by asking thousands of managers about their beliefs. Others try to measure the optimism (or pessimism) of a large number of consumers to get an idea of the situation on the demand side. All these methods have one thing in common - they are expensive. In order to construct a useful leading economic indicator, it is necessary to gather a huge amount of information that can only be collected by using surveys. These surveys are extremely costly to conduct because of the large sample sizes and the high periodicity, which are required.

The question arises whether the benefits from leading economic indicators based on these surveys justify their costs? Therefore, one should hope that these additional measures will have good predictive power.

The motivation for this diploma thesis is to analyse the performance of leading economic indicators in a European context. In order to do so, I will mostly focus on the Ifo Index, which seems to be the most noted and important leading indicator in Europe. This indicator is published by the well-known German Ifo Institute for Economic Research. The Ifo Index is based on approximately 7,000 monthly survey responses of firms in manufacturing, construction, wholesaling and retailing. The economists of the Ifo Institute explain the calculation of their index in the following way[1]. Using surveys, the firms are asked to give their assessments of the current business situation and their expectations for the next six months. They can characterise their situation as "good", "satisfactory" or "poor" and their business expectations for the next six months as "more favourable", "unchanged" or "more unfavourable". The balance value of the current business situation is the difference between the percentages of the responses "good" and "poor"; the balance value of the expectations is the difference between the percentages of the responses "more favourable" and "more unfavourable". The business climate is a transformed mean of the balances of the business situation and the expectations. For calculating the index values, the transformed balances are all normalised to the average of the year 2000. The so constructed indicator should give the condition and outlook of the supply side in the German economy.

In order to get a reference value, I will also analyse an indicator that is constructed to reflect the demand side. The so called Consumer Confidence is the most well-known leading indicator in this category. In Germany the prestigious GfK Group is assigned by the European Union to determine Consumer Confidence in monthly intervals[2]. This demand side indicator is designed to measure the degree of optimism of the economy that consumers are expressing through their activities of savings and spending. The procedure is similar to the Ifo index. By using surveys, thousands of households are asked about their attitudes and buying intentions. The saldo of positive and negative answers gives the Consumer Confidence. This measure has become an important indicator of the consumption component of the gross domestic product (GDP). Therefore Consumer Confidence is very useful for National Banks in order to determine the need for changes in the level of interest rates.

---

[1] See http://www.cesifo-group.de/portal/page/portal/ifoHome.
[2] See http://www.gfk.com/group/index.de.html

Figure 1.1. *Balance of positive and negative answers of Ifo Business expectations and Consumer Confidence for Germany.*
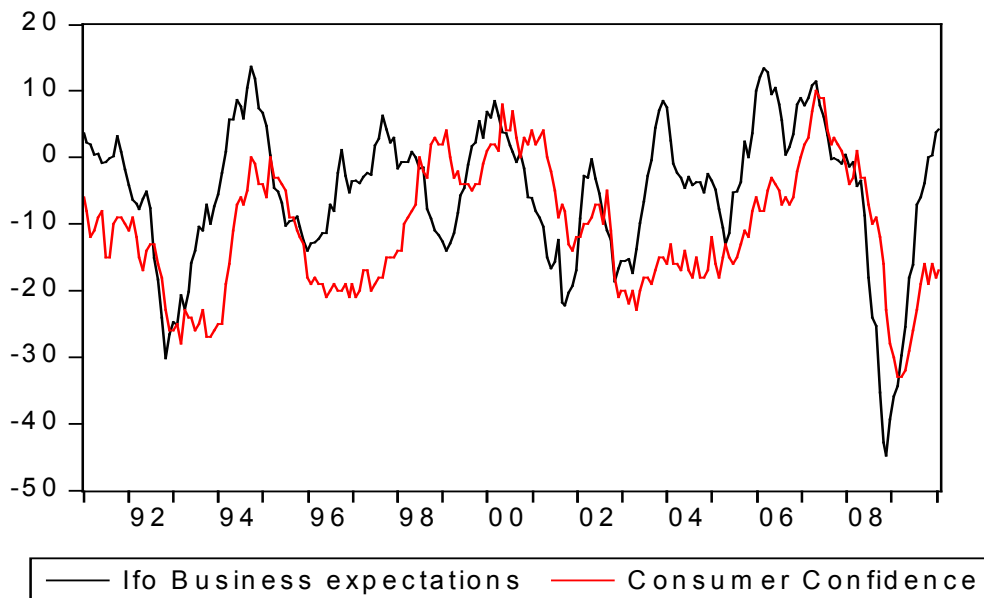
Beside these indicators, which are specially designed to predict future development in the economy, there are also some natural leading economic indicators, which seem to have good predictive quality. Stock market returns are an example of such a natural leading indicator. In the past, stock markets usually tended to decline before the entire economy slumped and began to improve before the economy recovered from a recession. For this reason stock indexes are often used as leading indicators which should reflect expectations for future profits. Therefore I will include the German stock index (DAX) in my analysis. The German stock index, established in 1987, is a so called blue chip index[3]. It contains thirty major German firms[4], which are part of the Frankfurt Stock exchange. DAX is said to measure the performance of these companies in terms of order book volume and market capitalization (Deutsche Börse[5]). The inclusion of the DAX in the analysis has a comparative impact. It should be interesting to see if there are differences in the predictive power of "natural" and specially designed leading indicators like Ifo Index and Consumer Confidence.

---

[3] A blue chip stock is a stock of a well-established company. These companies should have stable earnings and no extensive liabilities. Regular dividents are paid, even if the business is worse than usual.
[4] Such as Adidas, BMW, Lufthansa, SAP, Siemens, Volkswagen etc.
[5] see http://deutsche-boerse.com/dbag/dispatch/en/kir/gdb_navigation/home

Figure 1.2. *German stock index (DAX)*
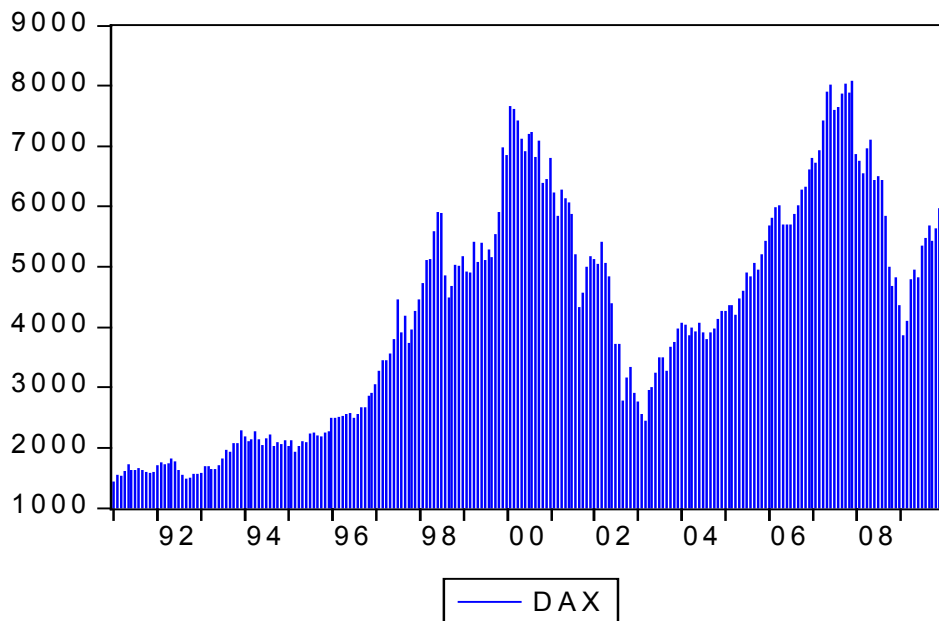
Another interesting feature is the question whether the Ifo index is also capable of predicting Austrian and European development. In the media Ifo Business expectations are often regarded as a European leading indicator, because of Germany's role as European driving force. Obviously Germany is by far the largest economy in the European Union and therefore it should be possible to extract some information about the future of the European economy from looking at the Ifo index. This should also be true for the development of the Austrian economy. Furthermore the strong foreign trade relations between Austria and Germany indicate a link between Ifo Business expectations and the future economic situation in Austria. For these reasons I will also cover the performance of Ifo Business expectations as a European and Austrian leading indicator in the thesis.

By using different testing procedures I will perform an analysis of the predictive power of different leading economic indicators. The diploma thesis is divided into two main parts. In the first one I will use time series models to investigate the causal relation of leading indicators and production. In order to do so, I will adopt the concept of Granger Causality and how this kind of causality can be tested within time series models.

After the time domain analysis I will bring in a completely different way of analysing the relation between the two processes. I will use Fourier analysis in order to decompose the data series in their basic frequency components, which allows a more sophisticated view on the performance of leading indicators. But before all this can be done the data that is used for the analysis has to be considered.

2. Data

2.1. Description

As already mentioned, the focus of this empirical work is on the potential predictive value of different leading indicators (Ifo index, Consumer Confidence Index, DAX) for German industrial production and the performance of the Ifo index as a European leading indicator. In order to accomplish this analysis, I will use time series models and spectral decomposition of economic variables. The used data series are publicly available and are provided by the European Federal Statistical Office (Eurostat)[6] and the Ifo Institute of Economic Research[7].

The Ifo Business expectation index and Consumer Confidence index are given by the balance between positive and negative answers, which should reflect the respondents' optimism or pessimism for the development of the German economy for the next six months[8]. The value for the German stock index is daily available and therefore had to be transformed to monthly data by simply using the value at the end of the month. Production account data are expressed as an index, with 2005 set as base index 1. I will use monthly data in the time span from 1991 to 2010, which leads to a sample size of 230 observations[9]. All time series[10] are already seasonally adjusted by the data providers.

---

[6] See http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database
[7] See http://www.cesifo-group.de/portal/page/portal/ifoHome
[8] See Chapter 1 for more information
[9] Austrian production index is only available from 1996 to 2010 due the new accounting rules by the European commission. Therefore the sample size reduces to 170 observations in the Austrian case.
[10] except the data for the German stock index (DAX), because there seems to be no seasonal pattern within the data.

2.2. Stationarity

Before starting with the empirical work, all time series have to be tested for stationarity. Stationarity is an important prerequisite for using time series models and spectral analysis. A series is stationary if the underlying stochastic process that generated the time series can be assumed invariant with respect to time. Therefore the time series must fulfil the following properties:

- constant mean
- constant variance and covariance
- short memory (i.e. rapidly declining ACF)

2.3. Testing for stationarity

It is always useful to look at the plot of the series. The visual observation can be seen as the first indicator to determine whether a series is stationary or not. In some cases it can be observed if the series shows signs of mean reverting behavior and constant variance. Furthermore a look at the Autocorrelation function (ACF) of the series will show if the prerequisite of short memory is fulfilled. The ACF of a stationary series is expected to die out rapidly.

After applying these methods to our data, it seems to be clear that the Ifo Business expectations are stationary. On the other hand the industrial production account data seems to be non-stationary. There is some evidence for a time trend and non-constant mean over the observed period of time.

For further review, I will use Dickey Fuller test procedure to determine in a more formal way, if the series are stationary. This test, introduced by Dickey and Fuller in 1979, is a so called unit root[11] test.

---

[11] A stochastic process has a unit root if 1 is a root of it's characteristic equation. If this is the case, the process will be non-stationary.

2.3.1. Dickey Fuller test (DF)

In order to run a DF test, we have to derive the following AR(1) model:

$$y_t = \rho y_{t-1} + \varepsilon_t \qquad\qquad (2.1)$$

by subtracting y_{t-1} from both sides of the model, we obtain:

$$y_t - y_{t-1} = (\rho - 1)y_{t-1} + \varepsilon_t \qquad\qquad (2.2)$$

or equivalently:

$$\Delta y_t = \beta_2 y_{t-1} + \varepsilon_t \qquad\qquad (2.3)$$

To test for the existence of a unit root, we have the following hypothesis:

$$H_0 : \beta_2 = 0 \quad \Rightarrow \rho = 1 \qquad\qquad (2.4)$$
$$H_1 : \beta_2 < 0 \quad \Rightarrow \rho < 1 \qquad\qquad (2.5)$$

where the null hypothesis is thus of a random walk (i.e. the series is integrated of order 1).

Under the null hypothesis y_t has a stochastic trend, but the non-stationarity condition of y_t implies that the OLS estimates of parameter ρ is biased and the t-test for β_2 is not standard t-distributed. In order to deal with the problem, Dickey/Fuller (1979) derived critical values for their test using Monte Carlo simulations[12]. This procedure has some critical influences on the test results.

---

[12] Monte Carlo simulations are a stochastic way to compute solutions for complex systems by using repeated random sampling. This method strongly relies on the law of large numbers.

These critical influences are the following:
- The computed critical values for the test are larger in their absolute value than those for the t-distribution. Therefore more evidence is needed to reject the unit root-hypothesis (null hypothesis).
- The critical values also highly depend on the sample size and also whether an intercept or a time trend is included in y_t.

After estimating

$$\Delta y_t = \beta_2 y_{t-1} + \varepsilon_t \tag{2.6}$$

we compare the t-statistic on β_2 with the critical values derived by the Dickey Fuller method. If the test statistic is less than the critical value, then the null hypothesis of a random walk can be rejected[13].

As mentioned before the estimation of the critical values are influenced by the existence of a time trend. Therefore alternative regressions should be taken into account:

$$\Delta y_t = \beta_1 + \beta_2 y_{t-1} + \varepsilon_t \tag{2.7}$$

$$\Delta y_t = \beta_1 + \beta_2 y_{t-1} + \beta_3 t + \varepsilon_t \tag{2.8}$$

Note that the distribution is effected by using equations with constants or/and trend components as regressors in the test construction[14].

Under the corresponding alternatives, these three equations[15] allow to test if y_t is

- A pure stationary process $\rightarrow$ (2.6)
- A stationary process with a constant $\rightarrow$ (2.7)
- A stationary process with a constant and a deterministic trend[16] $\rightarrow$ (2.8)

---

[13] This is usually a test on left tails, therefore it is non-symmetrical.
[14] This is the case in (2.7) and (2.8).
[15] i.e. equations (2.6), (2.7) and (2.8).
[16] such an equation can be used to test for trend stationarity, which is a weaker form of the concept of stationarity.

2.2.2. Augmented Dickey Fuller test (ADF)

So far we have only considered the case of an AR(1) process. Now we want to apply the Dickey Fuller framework to higher order autoregressive models. If we do so, DF testing procedure will lead to serial correlation in the residuals. Therefore DF test would not be valid. To overcome this problem, the Dickey Fuller regression should be extended by adding lags of the difference of y_t. The new equation has the form:

$$\Delta y_t = \beta_1 + \beta_2 y_{t-1} + \alpha_1 \Delta y_{t-1} + ... + \alpha_n \Delta y_{t-n} + \varepsilon_t \qquad (2.9)$$

Once again it should be mentioned that the equation can be extended by adding constant or a time trend depending on the problem.

The procedure to decide whether the null hypothesis can be rejected is the same as before. Again the coefficient on β_2 has to be compared to the critical values of the Dickey Fuller estimation. If the t-statistic on β_2 is smaller than the critical values, then the null hypothesis can be rejected.

2.2.3. ADF-test results

Table 2.1. shows the results of the Augmented Dickey Fuller tests (ADF) for the data in this empirical study. The appropriate number of lags used in the tests was determined by the Schwarz-Bayesian Information Criteria (SIC), such that the general models[17] were fitted and the SIC was minimized. The appropriate test equations were determined by looking at the plotted series.

---

[17] i.e. the model under the null hypothesis.

Table 2.1. *ADF – test results*

data series	test statistic	critical value at 5% level	prob.	stationary
German production				
ADF - test with trend and intercept	-2.892	-3.430	0.167	No
German production (log. diff.)				
ADF - test	-6.213	-1.942	0.000	Yes
Austrian production				
ADF - test with trend and intercept	-0.558	-3.437	0.980	No
Austrian production (log. diff.)				
ADF - test	-15.866	-1.943	0.000	Yes
Production in the euro area				
ADF - test with trend and intercept	-3.233	-3.430	0.081	No
Production in the euroarea (log. diff.)				
ADF - test	-4.171	-1.942	0.000	Yes
Ifo Business expectations				
ADF - test	-4.148	-1.942	0.000	Yes
Ifo Business expectations (first diff.)				
ADF - test	-7.038	-1.942	0.000	Yes
Consumer Confidence				
ADF - test with intercept	-2.127	-2.874	0.234	No
Consumer Confidence (first diff.)				
ADF - test	-13.821	-1.942	0.000	Yes
DAX				
ADF - test with intercept	-1.531	-2.874	0.516	No
DAX (first diff.)				
ADF - test	-13.917	-1.942	0.000	Yes

As already expected, the ADF-tests for the industrial production series strongly suggest non-stationarity. Because of this result I transformed the series by calculating the logarithmic first differences[18], which lead to monthly growth rates. The ADF-tests for the transformed series show that the null hypothesis of a random walk can be rejected at the 1% level, i.e. the series

---

[18] Logarithm of the first differences have been taken, because this is resulting to monthly growth rates, which is more common when using Production data. Therefore taking the logarithm should only simplify the interpretation of the tranformed series and will not affect the causality analysis.

are stationary. Therefore I will use the logarithmic differences of the production series for further analysis.

The ADF-tests for the leading indicators have different results. There is no transformation needed for Ifo business expectations index, because the time series is already stationary at the 1% level. On the other hand, the null hypothesis of a random walk cannot be rejected for Consumer Confidence and the German stock index (DAX). Both series seem to be non-stationary. This problem can be overcome by taking simple first differences. By doing so all series become stationary at the 1% level. In order to ensure homogeneity[19], I will also use first differences for Ifo Business expectations, although this series is already stationary in the raw case. The fact that all time series are equally transformed will simplify the interpretation of the upcoming results. After the transformations, all data sets are stationary with zero mean. Note that there should not be any difficulties with seasonality after deriving first differences and logarithmic differences, because all series are already seasonally adjusted by data providers. From now on I will refer to these transformed data series for further analysis.

2.3. Conclusions for chapter 2

I have just discussed the concept of stationarity which is an important prerequisite for using time series models and spectral decomposition in our framework. In order to get stationary series, the different time series were transformed in the following way:

Table 2.2. *Data transformation*

data series	transformation	variable name
German production	logarithmic first differences	dlprod_ger
Austrian production	logarithmic first differences	dlprod_aut
Production in the euroarea	logarithmic first differences	dlprod_eu
Ifo Business expectations	first differences	difo
Consumer Confidence	first differences	dcons
DAX	first differences	ddax

---

[19] i.e. all variables are transformed the same way.

3. Causality

3.1. Granger Causality

In my thesis I refer to the concept of Granger Causality (GC) introduced by Clive W. J. Granger (1969). His 1969 paper "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods" became a milestone for the study of causal relationship within the discipline of economic research. Although there have been a number of econometricians working on the topic of causality before (like Orcutt, Simon or Wiener), the concept developed by Granger is by far the most popular.

In 2003 Clive W.J. Granger was awarded with the Nobel Prize for Economic Science. The Royal Swedish Academy honoured him because of his fundamental discoveries in the analysis of time series. Granger's research had a huge impact on modern economics by changing the way economists analyse financial and macroeconomic data. His essays on the topic of causality are among his most important studies. Granger formulated his concept in a time when there was no causal measure, which was universally liked. His definition was elegant mathematically and also easy to implement for empirical research. Although there is still no complete consensus, Granger's contributions on causality and causality testing are at least the most accepted ones.

3.2. A Definition of Causality

For Granger the phrase "X causes Y" has to be handled with care, because the concept of causality is a subtle and difficult one. As mentioned before there was no universally accepted definition of causality. For this reason Granger tried to find a definition that would be reasonable for almost everyone. In his famous paper "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods (1969)" he suggested the following:

Let Ω_n represent all the information available in the universe at time n. Suppose that at time n optimum forecasts are made of Y_{n+1} using all the information in Ω_n and also using all of this information apart from the past and present values $X_{n-p}, p \geq 0$ of the series X_t. If the first forecast, using all the information, is superior to the second, than the series X_t has some

special information about series Y_t, not available elsewhere, and X_t is said to cause Y_t (Ashley/Granger/Schmalensee 1980).

Of course there has to be an assumption to determine which forecast is better than the other. Normally the criterion for a superior forecast is given by the comparison of the relative sizes of the forecast error variance. Therefore I will use the mean-square errors of the in-sample forecasts for this comparison.

It is also necessary to make some simplifications on the suggested definition of causation in order to apply it for empirical use. First the forecast-method will be restricted to only linear forecasts. I also have to replace the theoretical information set Ω_n by the past and present values of some set of time series, such that the information set is reduced to $\Omega_n : \{X_{n-p}, Y_{n-p}, Z_{n-p}, ..., p \geq 0\}$. It is clear that any causation will be relative to the set Ω_n, but it also important to mention that spurious results can occur if some relevant series is not in this set. For instance, if the information set Ω_n consists only of two different series X_t and Y_t, but there exists a third process Z_t, which is causing both X_t and Y_t within an enlarged information set Ω_n', then for the original set Ω_n, there may occur spurious causality between processes X_t and Y_t. This case is similar to the problem of spurious regressions arising from excluding a well needed explanatory variable in a statistical model.

In the context of this empirical work I can assume the simplest case of Ω_n consisting of just the values from series X_t and Y_t, because I will only test for causality in a bivariate framework.

With these simplifications the definition given above can be reduced to the following:

Definition 3.1.: Granger Causality

Let $MSE(Y)$ be the population mean-square of the one-step ahead forecast error of Y_{n+1} using the optimum linear forecast based on $Y_{n-p}, p \geq 0$ and let $MSE(X,Y)$ be the population mean-square of the one-step ahead forecast error of Y_{n+1} using the optimum linear forecast based on $X_{n-p}, Y_{n-p}, p \geq 0$. Then the process X causes process Y if $MSE(X,Y) < MSE(Y)$ (Ashley/Granger/Schmalensee 1980).

Granger Causality describes the extent to which a process Y is leading another process X. Therefore GC reflects a restricted sense of causality. If both X and Y are driven by a common third process with different lags, their measure of Granger causality could still be statistically significant. Moreover in our framework the phrase "X causes Y" does not really mean that Y is the result of X. Granger causality in the definition given above is only a measure for precendence and information content and therefore does not by itself indicate causality in the more common use of the term.

Definition 3.2.: Feedback

If $MSE(X,Y) < MSE(Y)$, and $MSE(Y,X) < MSE(X)$, we say that feedback is occuring. In the case of feedback not only X is causing process Y, but also Y is causing X. In other words there exists a causal relation from one process to the other and vice versa.

Definition 3.3.: Instantaneous Causality

Granger (1969) also introduced the so called Instantaneous Causality. This special case of causal relation is occuring if, in period t, adding x_{t+1} to the information set helps to improve the forecast of y_{t+1}. The concept of Instantaneous Causality is completely symmetric. That means that if there exists the so called Instantaneous Causality from process Y to X, there has to be Instantaneous Causality from X to Y.

Definition 3.4.: Causality Lag

Granger defined (1969) the (integer) causality lag m to be the least value of a given subsample, so that knowing the values Y_{n-j} with j = 0,1,…,m-1, will not improve the prediction of Y_t.

The definitions above have assumed that only stationary time series are used, because there are some difficulties arising in the non-stationary case. As defined in the foregoing chapter stationary processes are invariant over time. Therefore in the case of non-stationary series the mean-square errors of the post sample forecasts will depend on time t and therefore the existence of Granger causality would alter over time. Of course there could be some generalisations in this case (such as defining causality for a specified time t) in order to find

an operative solution for non-stationary time series. Granger (1969) mentioned that one could probably think of the existence of causality at a certain moment of time, but such a framework would be hard or even impossible to test in a statistical way. For this reason all data series used in this empirical thesis are transformed to be stationary (see chapter 2).

It can be argued that the use of the MSE as the criterion to measure the forecast precision is not the best one. It is obvious that this choice has a huge impact on the definition of causality. Using another measure may lead to different conclusions about the identification of causal relations between two series. To answer this criticism it should be mentioned that the concept of MSE seems to be the natural measurement in connection with linear forecasts. Moreover the MSE is easy to handle and simplifies the interpretation of the results. These advantages are good reasons for the use of mean squared forecast errors.

3.3. A short Introduction to Vector Autoregressions (VAR)

Before testing for Granger Causality in our data sets, I want to give a short introduction to Vector Autoregressions (VAR), because I will use this type of models for further causality analysis.

3.3.1. Vector Autoregression (VAR)

VAR models are the basic tool to analyse linear multiple time series generalizing the univariate AR models (Autoregressive models).

A multiple time series is nothing else than a vector of time series.

It consists of observations z_{kt} for variables $k = 1,...,K$ and for time points $t = 1,...,T$.

All the variables in a VAR are treated symmetrically by including for each variable an equation explaining its evolution based on its own lags and the lags of all the other variables in the model, i.e. each variable influences any other variable.

A vector autoregressive process is defined as a special case of a multivariate time-series process, such that z_t, a K-vector, depends on its past via the formula

$$z_t = v + A_1 z_{t-1} + ... + A_p z_{t-p} + u_t \qquad (3.1)$$

where v is a constant K-vector for the intercept and all A_j, $j = 1,...,p$ are $K \times K$-matrices, while u_t denotes a multivariate white noise process.

In this scheme, any of the component variables z_{kt}, $k = 1,...,K$ depends on p lags of itself and of the other K-1 component variables.

The individual equations are straightforward:

$$z_{mt} = v_m + \sum_{j=1}^{p}(A_j)_{m1} z_{1,t-j} + \sum_{j=1}^{p}(A_j)_{m2} z_{2,t-j} + ... + \sum_{j=1}^{p}(A_j)_{mK} z_{K,t-j} + \varepsilon_{mt} \qquad (3.2)$$

where $(A_j)_{lm}$ indicates the lm^{th} element of A_j.

3.3.2. Forecasting with VAR

A basic result in forecasting is that the expected squared error $E(\hat{y}_{t+h} - y_{t+h})^2$ is minimized among all predictors $\hat{y}_{t+h}$ – given the information set at time point t (Ω_t) – by using the conditional expectation

$$\hat{z}_{t+h} = E(z_{t+h}|\Omega_t) \qquad (3.3)$$

In a VAR model the conditional expectation looks as follows

$$E(z_{t+h}|\Omega_t) = v + A_1 z_1 + ... + A_p z_{t-p+1} \qquad (3.4)$$

with u_t assumed to be independent and therefore $E(u_{t+1}|\Omega_t) = 0$.

And similarly

$$E(z_{t+2}|\Omega_t) = v + A_1 E(z_{t+1}|\Omega_t) + A_2 z_t + ... + A_p z_{t-p+2} \tag{3.5}$$

The prediction error can be represented as a moving average

$$z_{t+h} - E_t(z_{t+h}) = \sum_{j=0}^{h-1} \Phi_j u_{t+h-j} \tag{3.6}$$

where Φ_j denotes the moving average matrices with the following form $\Phi_j = J A^j J'$.
The evaluation of the variance of the error can be done in this form

$$\Sigma_y(h) = \sum_{j=0}^{h-1} \Phi_j \Sigma_u \Phi'_j \tag{3.7}$$

Of course this is only a brief introduction to the method of Vector Autoregression. It should only give a short overview. For further information on VAR models, see Lütkepohl (New Introduction to Multiple Time Series Analysis, 2006) or Brockwell/Davis (Introduction to Time Series and Forecasting, 1996).

3.3.3. Model Specification

To find the right lag length j, it is mostly better to use more rather than fewer lags. The simple reason for this rule of thumb is that the theory of Granger Causality relies on the relevance of all past information. So one should think about the lag length j that corresponds to reasonable beliefs about the longest time over which process X could help to predict process Y. For instance, in the case of Ifo business expectations it is reasonable to use a lag length of six months ($j = 6$), because the Ifo Business Climate data reflect six months business expectations.

It also might be useful to look at the cross correlograms between dependant and explanatory variables. These show the cross correlation functions[20] and should give a first impression for choosing the lag order in our framework, as well as the visual inspection can give some indication of possibly existing relation.

As an example, Figure 3.1. shows the bivariate cross correlogram for German Industrial Production (logarithmic differences → dlprod_ger) and Ifo Business expectations (first differences → difo)[21].

Figure 3.1. *Cross correlogram for German Production and Ifo Business expectations.*
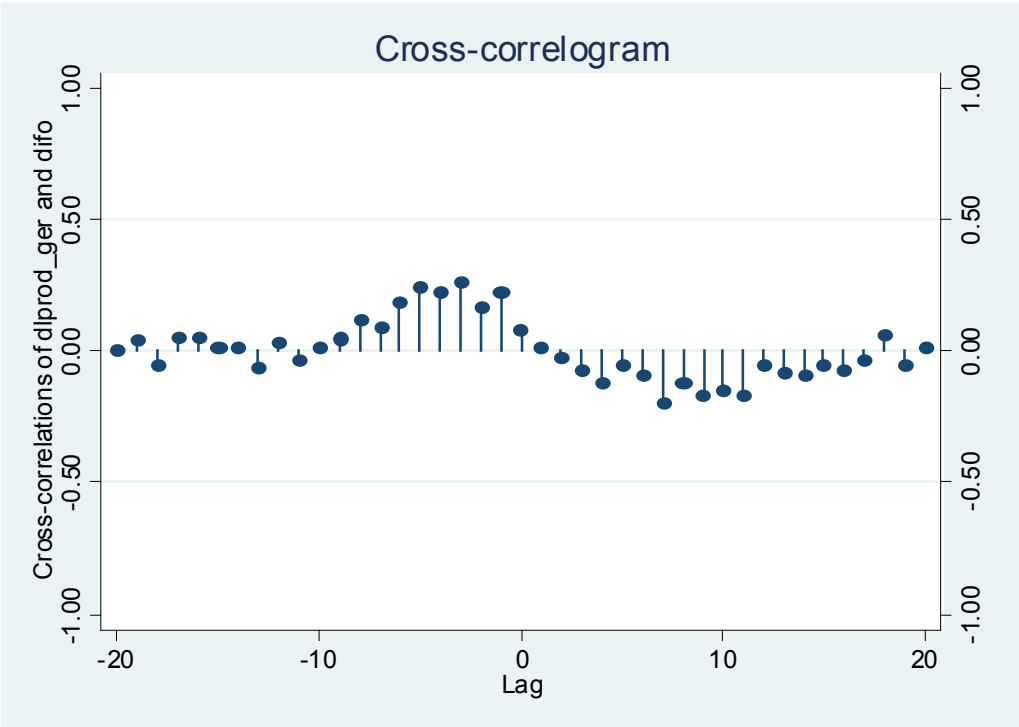
The cross correlogram indicates that the Ifo Business expectations are more of a short term indicator for German production. The correlation declines rapidly after the fourth lag. Note that the Ifo Index should reflect the management expectations for six months ahead. At lag six the correlation is already down at 0.18 (the highest value is at lag three). Therefore it can be

---

[20] The Cross correlation (function) is measures the linear dynamic interaction between two series. Consider two series x_i and y_i (with i = 0,1,2...N-1). The cross correlation r at delay d is defined as

$$r = \frac{\sum_i [(x_i - \bar{x}) \cdot (y_{i-d} - \bar{y})]}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_{i-d} - \bar{y})^2}}$$

[21] The plots for the other variables are available upon request.

said that the leading causal relation of the Ifo Business expectations might be shorter than expected. The result of rapidly declining cross correlations also holds for the other indicators (Consumer Confidence Index and DAX). The consumer confidence index seems to be equally linked to German production, although there is high correlation at $t+1$. This could imply that Consumer Confidence is probably also caused by changes of the production level. If this is true, there should be feedback between the two series. This can be determined by causality tests and will be done later in this chapter.

In the second part of my thesis I want to discuss the relevance of the Ifo Index for other European economies. To do so I estimated the cross correlation functions for Ifo Business expectations and the Austrian production, as well as for the Ifo Business expectations and the production within the Euro zone.

Figure 3.2. *Cross correlogram for Production in the Euro zone and Ifo Business expectations.*
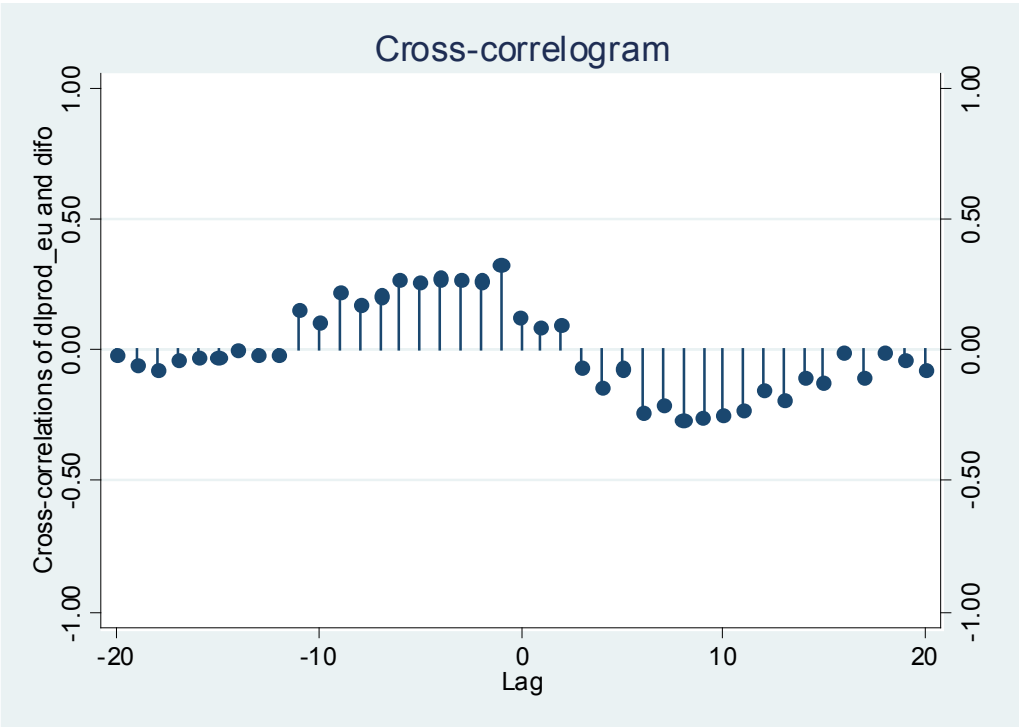
As observable in Figure 3.2., the cross correlogram of Production in the Euro area (logarithmic differences → dlprod_eu) and the Ifo Business expectations seems to be similar to cross correlogram in Figure 3.1.. The correlation is declining over time as seen before, but the speed of the decline is visibly different. The coefficient at lag six is as high as it is at lag

two. This might be interpreted as a longer lasting causal relation between the two series. The analysis of the cross correlogram for Ifo index and the Austrian production is also implying a longer relation as in the German production case, although the delay is not as long as in the case for the production in the euro zone.

It is also notable that uncontrolled correlation exists in some of the cross correlograms at certain higher lags/leads. These results are hard to interpret. Probably they are due the seasonal adjustment of the data or simple coincidental. Therefore I will not analyse these results any further.

In order to avoid overfitting the VAR model[22], it is also useful to take a look at certain information criteria like AIC (Akaike information), BIC (Schwarz information criterion) or HQ (Hannan-Quinn information criterion) for lag-selection. These criteria are statistics that measure the goodness of fit of an estimated statistical model. Given the data set, the model with the lowest information criteria value should be chosen. AIC, BIC and HQ have a similar form. They consist of two parts. The first is a standard goodness of fit value[23], which becomes smaller as the model becomes larger. Therefore there has to be a second part which penalizes the inclusion of additional lags. AIC, BIC and HQ differ in this penalty term. In general the AIC does not penalize larger models as strictly as the other two criteria. The milder penalty term results in a positive probability that the AIC asymptotically overestimates the VAR order, whereas the BIC and HQ criteria estimate the order consistently under quite general conditions if the actual data generation process (DGP) has a finite VAR order and the maximum tested order p_{max} is larger than the true order p (Lütkepohl 2004). On the other side the AIC tends to imply better model choice in smaller samples, so that the final selection of the information criteria depends on the specific circumstances.

In this thesis I refer to the concept of Granger causality and therefore I am more interested in the forecasting properties of the selected model. The final prediction error (FPE), introduced by Akaike (1969), is paying attention to this preference. The FPE is a calculation of the one step prediction mean squared error for a realisation of the process independent of the one observed (Lütkepohl 2004). When fitting different VAR models to the data, the maximum

---

[22] A VAR model is called to be overfitted, if there are too many lags included → i.e. an order higher than the optimal lag length is chosen.
[23] AIC, BIC and HQ use the maximum likelihood estimate of the error variance .

likelihood estimate of the error variance will decrease with increasing lag order p, while the estimation errors in the expanded set of fitted parameters will increase the final prediction error. Following this procedure, the lag order p with the lowest FPE value has to be chosen. Note that FPE and AIC have the same large-sample and similar small-sample properties. As already mentioned, HQ and BIC have stronger penalty terms[24]. Thus, the decision for the FPE essentially is an AIC decision.

Because of the proximity to the method in this thesis, I will refer to the FPE criterion, if the different information criteria do not provide a clear solution for finding the order of the respective models.

Table 3.1. Lag Length criteria

VAR model	AIC	BIC	HQ	FPE
model 1: German Production, Ifo Index	3 lags	3 lags	3 lags	3 lags
model 2: German Production, Consumer Confidence	3 lags	1 lag	3 lags	3 lags
model 3: German Production, DAX	3 lags	1 lag	1 lag	3 lags
model 4: Austrian Production, Ifo Index	4 lags	1 lag	2 lags	4 lags
model 5: Production in the euro area, Ifo Index	9 lags	3 lags	3 lags	9 lags

The results of the lag selection criteria are shown in the table above. Note, that all series are transformed as described in chapter 2. The maximum lag order p_{max} was set to 8 lags for the first three models and 12 lags for model five, because of the longer lasting correlation observed in the cross correlogram. Most of the results are straightforward – for model one and model two there is strong evidence for a VAR(3) process. The suggestions for the other three models are not so obvious. The information criteria differ for model three, four and five. As mentioned before, I use the final prediction error (FPE) in such a situation. Therefore I will take a VAR(3) for model three, a VAR(4) for model four and a VAR(9) for model five.

After determining the lag order p, I estimated the respective (unrestricted) VAR models[25], which will be used for the causality analysis in the time domain. The coefficient structure for all models has been checked for their standard properties. It could be determined that all defined VAR processes are stable.

---

[24] BIC has the strongest penalty term.
[25] The complete estimation results (such as coefficients etc.) are available upon request.

3.4. Testing for Granger Causality in the time domain

There are different ways of testing for causal relations mostly depending on the exact definition of causality. A very common and simple test procedure for Granger Causality in the time domain is the so called Granger-Wald test. As already mentioned, Granger (1969) argued in his definition that it has to be determined how much of the current process Y is explained by its own past values and then whether adding another lagged value of X can improve the explanation. Process Y is Granger caused by process X if the addition of X improves the prediction of Y, or equivalently if the coefficients on the lagged X's are zero or not.

3.4.1. The Granger-Wald test

One advantage of the concept of Granger Causality is the easy implementation within a VAR model. To show that, suppose that the K-variate process z_t (see 3.3.1.) consists of only two components y_t and x_t , with $z_t' = (y_t', x_t')$. Obviously we are now looking at a bivariate process. Further, suppose that this process has a moving average representation[26] of the following form:

$$z_t = \mu + \sum_{j=0}^{\infty} \Phi_j u_{t-j} = \mu + \Phi(L)u_t \qquad \text{with } \Phi_0 = I_k \tag{3.8}$$

Then, the infinite MA model can be written in its partitioned form:

$$z_t = \begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \Phi_{11}(L) & \Phi_{12}(L) \\ \Phi_{21}(L) & \Phi_{22}(L) \end{bmatrix} \cdot \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \tag{3.9}$$

Now that we have defined Granger Causality and the partitioned form in an estimated VAR system (with MA representation), we can test for zero constraints of the estimated coefficients. It can easily be shown that X does not Granger Cause Y if and only if $\Phi_{12} \equiv 0$, i.e. if $\Phi_{12,j} = 0 \qquad \forall j = 1,...,\infty$.

---

[26] In the moving average (MA) representation of the process, z_t is expressed in terms of past and present error or innovation vectors U_t and the mean term μ . For more details see Lütkepohl (2006).

This result is more important for theoretical reasons and the implementation of other VAR methods (like Impulse Response Analysis), nevertheless it should be mentioned, because, based on this result, more common versions in the VAR context have been constructed.

For applied purposes it is more important to know how a Granger-causal relation can be expressed in VAR models. Luckily, there exists a completely identical property for all finite order VAR systems, such as in the partitioned VAR representation:

$$\begin{bmatrix} A_{11}(L) & A_{12}(L) \\ A_{21}(L) & A_{22}(L) \end{bmatrix} \cdot \begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \tag{3.10}$$

$\Rightarrow$ x does not Granger cause y if and only if the operator $A_{12} \equiv 0$. Vice versa, Y does not Granger cause X (Feedback) if and only if $A_{21} \equiv 0$ [27].

Therefore we have to test the null hypothesis

$$H_0 : A_{12} = 0 \tag{3.11}$$

against the alternate hypothesis

$$H_1 : A_{12} \neq 0 \tag{3.12}$$

Under the null-hypothesis x does not Granger cause y. Clearly under the alternate-hypothesis, there is evidence for Granger Causality.

Note that (if feedback is not of interest) the second set of equations is completely irrelevant. For the Granger-Wald Causality test only the restricted equations are used.

The hypothesis given above can be tested using different testing procedures. The most common are χ^2- and F-tests based on the Wald principle.

Unfortunately, there are some problems arising from using these standard tests in our framework. Lütkepohl (2004) mentioned that χ^2- and F-tests may have non-standard asymptotic properties if the estimated VAR model consists of I(1) processes. As described in the previous chapter, I transformed the data to first and logarithmic first differences in order to get stationary time series. These transformations may influence the testing results.

---

[27] The proof of equivalence uses the facts that the inverse of a block triangular matrix is again block triangular and that the leading matrix $A_0 = I_K$.

Toda/Phillips (1993) showed that Wald tests for Granger causality can result in non-standard limiting distributions depending on the cointegration properties[28] of the variables.

Therefore Toda/Yamamoto (1995) suggest a more specific way to test for Granger causality. The issue is that there exists a singularity of the asymptotic distribution of the estimators, which lead to the non-standard asymptotic properties of the standard tests on the coefficients of cointegrated VAR processes (Lütkepohl 2004). Fortunately, this problem can be solved without much of an effort. The singularity can be removed by simple overfitting the VAR model. In other words, we estimate the VAR model of the order p+1 instead of the real order p. In doing so we ensure that the relevant parameters have a non-singular asymptotic distribution. As a consequence, the shortcomings of the standard tests are not a problem if the tests are based on the estimated parameters of an overfitted model, where the zero restrictions are only performed on the relevant parameters and the extra parameters are ignored. This procedure is universally applicable. It is not necessary to know the cointegration properties of the system in a great detail.

After estimating the overfitted VAR(p+1) model, I will only use the first p coefficient matrices for testing the zero restrictions, which lead to the following null hypothesis:

$$H_0 : A_i = 0 \quad \text{with } i = 1,..., p \tag{3.13}$$

Note, that although the test is based on a VAR(p+1) model, the augmented lags (p+1) are not relevant. The only reason for overfitting the model is to ensure the standard asymptotic properties.

The Wald statistic has the usual $\chi^2(p)$ – distribution and can be used for the Granger causality test. Lütkepohl (2004) mentioned that it is advisable to use an F-version of the test, because of its better approximation of the desired size of the test. The denominator degrees of freedom are obtained as the total number of observations used for the estimation of the model (2T) minus the total number of estimated parameters.

---

[28] Two time series are said to be cointegrated if they have a common stochastic trend. If there is cointegration in the data sets, the VAR method may not be the proper technique for analysis. In such a case it might be better to use other types of models, which support the analysis of the cointegration structure.

3.4.2. Testing for instantaneous causality

So far we have only talked about testing for Granger causality and did not discuss possible testing procedures for instantaneous causality. Tests for this kind of causal relation can be done similarly to the testing procedure already explained, because instantaneous causality can be expressed in terms of zero restrictions for the error terms. More precisely one has to determine if instantaneous residual correlation exists. This procedure does not lead to problems of non-standard properties like in the Granger Causality framework, because the asymptotic properties of the estimator of the residual covariance matrix of a VAR process are unaffected by the degree of integration and cointegration in the variables (Lütkepohl 2004). Therefore, under the standard assumptions, the estimated test statistic based on the already discussed Wald principle has an asymptotic χ^2 - distribution.

Hence, it is not necessary to estimate an overfitted model. The test for instantaneous causality is only based on the residuals of an VAR(p) model, where p is the optimal lag length. Unlike the Granger-Wald test, I will use the χ^2 - approximation for the instantaneous causality test, instead of the F-version. The number of degrees of freedom of the approximating χ^2 - distribution is equal to one, because only one correlation coefficient has to be tested to be equal to zero. As already mentioned earlier, the concept of instantaneous causality is completely symmetric. Therefore the direction of the instantaneous causal relation does not matter for the test result. Both directions are identical. If there is evidence for instantaneous causality from X to Y, then there exists the same relation from Y to X.

3.5. Test results in the time domain

The results of the previously explained causality tests are given in tables below.

The first one provides the results for German leading indicators and German production. My analysis includes the Granger-Wald test for Granger Causality (GC), as well as testing for feedback and instantaneous causality. The estimated test statistics and p-values can be used to compare the performance of the different indicators in the time domain. I suggest that higher test statistics and the resulting lower p-values give more evidence for rejecting the null hypothesis of no existing Granger causality. Therefore I can compare the performance of the leading indicators by comparing their p-values. Note that the data is transformed as explained in part 2 – for instance the null hypothesis "Ifo Index does not cause German production" actually means that a change in Ifo business expectations does not cause a change in German production.

Table 3.2. *test results for German leading indicators.*

Null hypothesis	Granger-Wald test statistic	p-value	GC	Feedback	instant. causality
Ifo Index does not cause German Production	8.312	0.000	Yes***	No	No
Consumer Confidence does not cause German Production	2.876	0.036	Yes**	Yes***	Yes***
DAX does not cause German Production	4.728	0.003	Yes***	No	Yes**

*,**,*** imply significance at the 10-, 5- and 1 % level

A look at the results shows that there is evidence for Granger causality for each leading indicator and German production. The causal impact of the Ifo Index and the DAX for German production is significant at the 1% level, with Ifo Index having the highest test statistic at 8.3. These two indicators seem to have good statistical properties to predict the upcoming German production.

The Consumer Confidence does not have the same properties in terms of statistical significance. Unlike the Ifo Business expectations and the DAX there is evidence for Granger causality only on the 5% level. Of course this result also implies good prediction properties for German production, but compared with the other two indicators, the Consumer Confidence Index has to be ranked last. Furthermore it is problematic that there is feedback

occurring between consumer confidence and German production. This result is in line with the findings in part 3.3.3., where the cross correlogram implied high correlation at $t+1$. With existing feedback, it is not possible to distinguish between effect and cause variable. This is hard to take in our framework, because the causal direction for an appropriate leading indicator should be clear. The test result implicates that Consumer Confidence is also affected by changing production levels. Note that the feedback effect is highly significant at the 1% level, while the causal link from Consumer Confidence to German Production is only statistically significant at the 5% level. Therefore German production might be a better leading indicator for Consumer Confidence, not vice versa. It is also notable that there is instantaneous residual correlation between consumer confidence and German production, as well as for the German stock index and German production. The resulting instantaneous causality in these cases implies that there is an immediate impact from one series to the other.

Table 3.3. The Ifo Index as an European indicator.

Null hypothesis	Granger-Wald test statistic	p-value	GC	Feedback	instant. causality
Ifo Index does not cause German Production	8.312	0.000	Yes***	No	No
Ifo Index does not cause Austrian Production	2.687	0.032	Yes**	Yes**	No
Ifo Index does not cause Production in the euro area	3.592	0.000	Yes***	Yes***	No

*,**,*** imply significance at the 10-, 5- and 1 % level

The table above provides the test results for the analysis of the performance of the Ifo Index as a leading indicator for European and Austrian production. We already witnessed a longer lasting causal link in this context. Remember that the correlograms showed prolonged correlation structure and that the chosen model types have been of a higher order than for the analysis for German production.

The Granger-Wald tests give evidence for existing Granger causality. The Ifo Business expectations cause the production in the euro area at the 1% level. The significance level in the Austrian case is not so high, but there exists evidence for causality at least at the 5% level. These results suggest that the Ifo Business expectation Index seems to be an appropriate measure for the dynamics of the European production.

Beside these positive results, there are also some issues arising when applying the Ifo Index to European and Austrian production. In both models there is strong evidence for feedback. As discussed before in the case of consumer confidence, occurring feedback make it hard to distinguish between effect and cause. This scenario is not very convenient in our framework, because an efficient leading indicator is expected to cause the depending variable and not vice versa. Obviously the existence of feedback does not have a direct negative impact on the predictibility of the leading indicator in Granger's concept, but in a more common sense of causality it is desireable to know the exact causal chain. Therefore these feedback mechanisms should be taken into account when using the Ifo business expectations index as an indicator for other countries than Germany[29].

3.6. Conclusions for chapter 3

The following conclusions have been drawn in the previous chapter:

- Ifo Business expectations index and DAX seem to be viable leading indicators for German production.

- In terms of statistical properties the Ifo business expectation index can be ranked as the best leading indicator for German production.

- There is evidence that consumer confidence is affected by German production, and therefore consumer confidence might not be a viable leading indicator in a more common sense of the term.

- The Ifo business expectations also cause the Austrian production and the production in the euro zone, although the existing feedback in these cases should be taken into account.

---

[29] As well as applying Consumer Confidence on German production.

4. Spectral Methods

4.1. Introduction

So far I have used cross correlograms and causality tests in the time domain to analyse the performance of leading indicators in predicting the upcoming economic development. In the next step I will also use spectral methods to find lead-lag relations among our data sets. The origin of these methods is the so called Fourier analysis, which is named after the French mathematician Jean Baptiste Joseph Fourier (1768-1830), who tried to find rules when it was possible to write general functions by sums of simpler trigonometric functions[30]. In other words, Fourier analysis is a method to approximate functions by breaking them into their basic components, which should be easier to understand. These components are called Fourier series, which consist of sines and cosines (or complex exponentials).

The spectral representation of a stationary time series X_t essentially decomposes X_t into a sum of sinusoidal components with uncorrelated random effects (Brockwell/Davis 1991). The idea of decomposing economic time series allows a more detailed analysis of the correlation between them. It is possible to find correlations in specific frequency bands, which cannot be found in the raw data. The analysis of the spectral representation is also called the frequency domain analysis or spectral analysis. It is equivalent to the time domain approach based on the covariance function. The difference is that the time domain approach shows how a series changes over time, while the frequency domain approach indicates how much of the series lies within each given frequency band over a range of frequencies. The advantage of the frequency domain is that it gives a completely different way of viewing the process of interest. In the following I will discuss the estimation of the spectral representation and how it can be used to give a more insightful view on the structure of the underlying process. It is also necessary to give a brief introduction to filtering and smoothing techniques. Finally I will extend the framework to the bivariate case by using cross spectral methods. These methods will be used to analyse the causality (in the frequency domain) between economic leading indicators and production like in the previous chapter.

---

[30] In mathematics, the trigonometric functions are functions of an angle. Also known as sine, cosine and tangent.

4.2. The spectral estimation of an economic time series

Economic theory suggests the possibility to decompose economic time series into three different main parts: long, medium and short run behaviour. These three main parts correspond to different kinds of movements. Usually in the case of an economic time series there exists a slowly evolving, often linear movement, a faster moving part and a rapidly varying and temporary component. These parts can be understood as the trend (long run), the business cycle (medium run) and the seasonality (short run). Of course this separation of the components is a very theoretical one. In reality it is not easy to distinguish between these parts, because of the uncertainty about the underlying data generating process.

In modern econometrics there exist various techniques to isolate each of the mentioned characteristics within a time series. A number of smoothing methods have been introduced in order to extract business cycles. Simple methods like moving averages are often used to eliminate the rapid oscillating components, first differences can be taken to exclude the slowly evolving, long term components. Although all these techniques are not wrong in theory, they cannot be used to formally decompose the data series (Iacubucci 2003). One way to overcome these shortcomings is to use Fourier analysis in order to perform the separation of a signal[31] into different periodic parts, which can be expressed as frequencies driven by sines and cosines. This might be seen as a system responding to different driving frequencies by estimating linear combinations of sine and cosine functions. In such an imagination the time domain approach may be thought as a regression of the present on the past, whereas the frequency domain approach may be considered as a regression of the data on periodic sines and cosines (Shumway/Stoffer 2006).

The following theoretical part is mainly based on Shumway/Stoffer (2006), who provide an excellent introduction to the application of spectral methods.

---

[31] In our case the term signal means the time series, which is used for analysis.

4.2.1. Frequency measures

Before we can discuss the process of transforming a time series to the frequency domain[32] more precisely, we have to identify the dominant frequencies in our data series. This objective is a fundamental task of spectral analysis. In many cases there will be a lot of different frequencies, which coexist. In order to keep the framework as simple as possible, I will use cycles per data point[33] as the frequency measure (ω) and discuss the implication of certain frequencies in the context of the different problems. For $\omega=1$ the time series shows a cycle per time unit, $\omega=0.5$ features a cycle within two time units, for $\omega=0.1$ there is a cycle every tenth time unit, and so on. Obviously there have to be at least two data points to identify a circle. For this reason the frequency $\omega=0.5$ is the highest frequency of interest and is called the folding frequency. This is the highest frequency, which can be observed in discrete sampling. Because of this limitation of the signal[34], all higher frequencies will appear in lower frequencies. This effect is called aliasing.

As an example for analysing the causal relation of leading economic indicators and production, we may assume a predominant frequency of one cycle per year. Because of the use of monthly data, this is corresponding to one cycle every twelve months, or 0.083 cycles per observation. Another important measure is the period of the time series (T), which is defined as the number of data points in one cycle:

$$T = \frac{1}{\omega} \tag{4.1}$$

Therefore the predominant period in our example is twelve months per cycle. This general definition of periodicity can be used in order to get a more precise notion by adding some terminology. To find the rate at which a time series oscillates, we can first define a cycle as one period of a sine or cosine function. This function is defined over the interval of length 2π.

Now that we have introduced frequencies and the period of the time series, we can give a more sophisticated view on Fourier analysis and the spectral estimation methods.

---

[32] i.e. the linear transformations, which matches sines and cosines of different frequencies against the underlying data.

[33] Alternatively it is common to use the frequency measure $\lambda = 2\pi\omega$ that would give radians per data point. This measure is more common in statistics.

[34] i.e. sampling with some finite time period.

4.2.2. The Spectral Density

Before applying spectral methods, we should think about the characteristics of these methods. The idea of decomposing a time series into different periodic components is fundamental in the spectral representation. Certainly the first problem that has to be covered is the applicability of this decomposition. Mathematically it can be proved that any stationary time series may be thought of as the random superposition of sines and cosines oscillating at various frequencies (Shumway/Stoffer 2006). In other words, spectral methods can be applied to every stationary time series. The next question of interest is whether there is any meaningful spectral representation for the autocovariance function in the time domain for each time series? This question can be answered positively – there exists such a representation. It is called the spectral density, which essentially measures the variance in a certain kind of periodic oscillation.

The spectral density is a positive real function of a frequency variable (ω), which is linked with a stationary stochastic process. To help understanding the procedure more intuitively, one can see the spectral density as a function that captures the frequency content of a process and helps to find underlying periodicities.

First, consider a series X_t. This series is a stationary time series with zero mean[35] and an autocovariance function $\gamma(h) = E\left[(x_{t+h} - \mu)(x_t - \mu)\right]$. Under these assumptions there exists the so called spectral distribution function $F(\omega)$, which is unique monotonically increasing, such that the autocovariance function can be rewritten as

$$\gamma(h) = \int_{-0.5}^{0.5} e^{2\pi i \omega h} dF(\omega) \tag{4.2}$$

where $e^{i\omega} = \cos(\omega) + i \cdot \sin(\omega)$ and $i = \sqrt{-1}$. Because of the absolute summability of the autocovariance function, it can be shown that the spectral distribution function is continuous and therefore there is a function $f(\omega)$, such that

$$dF(\omega) = f(\omega)d\omega \tag{4.3}$$

---

[35] This prerequisite is fulfilled for all transformed time series used in this thesis.

This leads to the following property. If the autocovariance function $\gamma(\cdot)$ satisfies the condition $\sum_{h=-\infty}^{\infty}|\gamma(h)| < \infty$, then it can be written as

$$\gamma(h) = \int_{-0.5}^{0.5} e^{2\pi i \omega h} f(\omega) d\omega \tag{4.4}$$

Under these assumptions the so called spectral density (or spectrum) of X_t at frequency ω is defined by

$$f(\omega) = \sum_{-\infty}^{\infty} e^{-2\pi i \omega h} \gamma(h) \qquad \text{with } -0.5 < \omega < 0.5 \tag{4.5}$$

The interval of interest can be limited to –0.5 and 0.5, which represent the folding frequency mentioned in the previous part (4.2.1.). The so defined spectral density is the spectral counterpart of the autocovariance function in the time domain. In other words, the autocovariance function $\gamma(h)$ and the spectral density function $f(\omega)$ present the same information content, but expressed in a different manner. While the autocovariance function is expressed in terms of time lags, the spectral density shows the same information in terms of different cycles.

For the fact that the covariance function $\gamma(h)$ is non-negative definite, we get the following basic properties of the spectral density $f(\omega)$:

-   $f(\omega) \geq 0 \qquad \forall \omega$

-   $f(\omega)$ is an even function[36]

-   $\gamma(0) = \text{var}(x_t) = \int_{-0.5}^{0.5} f(\omega) d\omega$, which shows the total variance as the integrated

    spectral density function $f(\omega)$ over all feasible frequencies.

---

[36] i.e. $f(\omega) = f(-\omega)$

The formulas given for the autocovariance function in (4.4) and for the spectral density function in (4.5) are also called Fourier transform pairs. Such pairs are defined in the following way:

Given a sequence of real numbers $\{a_t; t = 0, \pm 1, \pm 2, ...\}$, which fulfils the absolute summability condition, i.e.

$$\sum_{t=-\infty}^{\infty} |a_t| < \infty, \tag{4.6}$$

a general Fourier transform pair can be defined as:

$$A(\omega) = \sum_{t=-\infty}^{\infty} a_t e^{-2\pi i \omega t} \tag{4.7}$$

and

$$a_t = \int_{-0.5}^{0.5} A(\omega) e^{2\pi i \omega t} d\omega \tag{4.8}$$

The Fourier transform pairs given in (4.4) and (4.5) are major parts in analysing discrete time series in the spectral domain. It can be shown that these Fourier transform pairs exist and are unique if the summability condition $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$ is fulfilled.

4.2.3. The Periodogram

The periodogram is the sample-based counterpart to the population based spectral density. Before explaining this concept more precisely, we have to define the discrete Fourier transform (DFT):

Given a finite time series X_t with $x_1, x_2, ..., x_n$, the discrete Fourier transform (DFT) can be expressed by

$$d(\omega_j) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} x_t e^{-2\pi i \omega_j t} \qquad \text{with } j = 0, 1, ..., n-1 \tag{4.9}$$

The frequencies $\omega_j = \frac{j}{n}$ are known as the Fourier frequencies or fundamental frequencies. The inverse DFT can be estimated by simple linear transformation such that

$$x_t = \frac{1}{\sqrt{n}} \sum_{j=1}^{n-1} d(\omega_j) e^{2\pi i \omega_j t} \qquad \text{with } t = 1, 2, ..., n \tag{4.10}$$

The discrete Fourier transform (DFT) decomposes the data series into components of different frequencies. For a large enough sample size n, there exists the fast Fourier transform (FFT)[37], which is an efficient algorithm to estimate the discrete Fourier transform (DFT) or its inverse and is used in the most statistical software packages, which provide spectral estimation.

With the definitions given above, we can consider the periodogram as the squared modulus of the discrete Fourier transform given in (4.9):

$$I(\omega_j) = \left| d(\omega_j) \right|^2 \qquad \text{with } j = 0, 1, ..., n-1 \tag{4.11}$$

The periodogram can be seen as some kind of sample density function of process X_t.

---

[37] introduced in 1965 by Cooley/Tukey.

It is often useful to break down the discrete Fourier transform into real and imaginary parts. Therefore the next two transforms are presented:

Like before, a finite time series X_t is given. Then the cosine transform is given by

$$d_c(\omega_j) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} x_t \cos(2\pi\omega_j t) \tag{4.12}$$

and analogue the sine transform is given by

$$d_s(\omega_j) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} x_t \sin(2\pi\omega_j t) \tag{4.13}$$

where $\omega_j = \dfrac{j}{n}$ for $j = 0,1,...,n-1$ are the Fourier frequencies.

Because the discrete Fourier transform (DFT) can be written in terms of the sine and cosine transforms, such that

$$d(\omega_j) = d_c(\omega_j) - id_s(\omega_j), \tag{4.14}$$

the periodogram can be expressed as

$$I(\omega_j) = d_c^2(\omega_j) + d_s^2(\omega_j). \tag{4.15}$$

4.2.4. The smoothed periodogram

In order to complete the discussion on univariate spectral methods, we have to introduce nonparametric spectral estimation for a more applied use, because the periodogram explained in the previous chapter is only asymptotically unbiased for the theoretical power spectrum. In our case of finite time series, the formula given in (4.11) is problematic since the variance at a given frequency does not decrease as the number of observations used in the estimation increases. Because of its instability the spectral estimation results are difficult to interpret. Fortunately there exists a more stable estimator[38] without much of an extra effort. With the technique of windowing it is possible to smooth all abrupt variations and to minimize spurious fluctuations. The result of the windowing is the smoothed spectrum. In order to explain this concept we apply the so called frequency band B, which consists of a number of L fundamental frequencies[39] situated around the Fourier frequency $\omega_j = j/n$, that are near to the frequency of interest, ω, such that

$$B = \left\{ \omega : \omega_j - \frac{m}{n} \leq \omega \leq \omega_j + \frac{m}{n} \right\} \tag{4.16}$$

Note that L has to be an odd number (i.e. $L = 2m+1$), such that the spectral values[40] in B are approximately equal to $f(\omega)$. This definition can be approached with relatively large sample sizes and has the advantage of relatively constant spectral values.

When using the defined frequency band, we need a new estimator for the periodogram. Instead of the concept introduced in the previous chapter, we use the smoothed periodogram $\hat{f}(\omega)$ as the average of the periodogram values over the frequency band B:

$$\hat{f}(\omega) = L^{-1} \sum_{k=-m}^{m} I(\omega_j + k/n) \tag{4.17}$$

It can be proved that the smoothed periodogram is equivalent to splitting the time series into different sub-series with the same length, estimating their spectra and then taking their mean.

---

[38] i.e. the estimator has a smaller variance.

[39] Where L has to be strictly smaller than n.

[40] The spectral values in the frequency band are given by $f(\omega_j + k/n)$, with $k = -m,...,0,...,m$.

For a more visual interpretation one can think of the smoothed spectrum as the periodogram seen through a window opened on an appropriate interval around the Fourier frequency. The width of the window can be chosen by simple testing different sizes, i.e. starting at a relatively small value of L, and then widening the window until spectral stability is obtained, which means that the estimation remains nearly unchanged for higher values. This procedure is also known as window-closing. The choice of the frequency band is a rather difficult one, because it has to be large enough to let all the fundamental details of the spectrum appear, but not too large in order to prevent the generation of spurious peaks (Iacobucci 2003). Therefore it is necessary to find a tradeoff between accuracy and stability of the spectral estimator.

4.3. Cross-Spectral methods

So far we only discussed univariate spectral techniques, which allow the identification of movements inside each series. The investigation of the relation of two different time series in the frequency is obviously linked to these concepts introduced in the previous chapters. For the study of causality between production and leading indicators, we refer to the bivariate extension of spectral analysis, i.e. cross-spectral analysis, which allows to describe pairs of stationary time series (x_t, y_t) in the frequency domain, by decomposing their covariance functions into frequency components. Therefore the difference between univariate and bivariate methods is simply substituting the covariance function for the autocovariance function. The covariance function of x_t and y_t is given by:

$$\gamma_{xy}(h) = E\left[(x_{t+h} - \mu_x)(y_{t+h} - \mu_y)\right]$$
(4.18)

with μ_x and μ_y being the sample means of series X_t and Y_t.

The covariance function in (4.18) has the representation

$$\gamma_{xy}(h) = \int_{-0.5}^{0.5} f_{xy}(\omega)e^{2\pi i\omega h}d\omega \qquad \text{with } h = 0,\pm 1,\pm 2,..., \tag{4.19}$$

and the cross-spectrum is given by the Fourier transform:

$$f_{xy}(\omega) = \sum_{h=-\infty}^{\infty}\gamma_{xy}(h)e^{-2\pi i\omega h} \qquad \text{with } -0.5 \le \omega \le 0.5 \tag{4.20}$$

When assuming that the absolute summability condition for the autocovariance function is also applicable for the covariance function, the cross-spectrum is in general a complex function[41] and can be written as:

$$f_{xy}(\omega) = c_{xy}(\omega) - iq_{xy}(\omega) \tag{4.21}$$

with

$$c_{xy}(\omega) = \sum_{h=-\infty}^{\infty}\gamma_{xy}(h)\cos(2\pi\omega h) \tag{4.22}$$

and

$$q_{xy}(\omega) = \sum_{h=-\infty}^{\infty}\gamma_{xy}(h)\sin(2\pi\omega h) \tag{4.23}$$

The real part $c_{xy}(\omega)$ is also called the cospectrum, while the imaginary part $q_{xy}(\omega)$ is also known as the quadrature spectrum.

The most important application of the cross-spectrum is the prediction of an output series Y_t from a certain input series X_t in the frequency domain. This is exactly what we are looking for in our framework and what we already investigated in the time domain[42].

---

[41] i.e. a complex-valued function, since the autocovariance function is not even.
[42] see chapter 3.

The measure for the strength of the relation of two time series in the frequency domain is given by the squared coherence function:

$$\rho_{xy}^2(\omega) = \frac{\left| f_{xy}(\omega) \right|^2}{f_{xx}(\omega) f_{yy}(\omega)} \qquad (4.24)$$

where $f_{yx}(\omega)$ is the cross spectrum as given in (4.20)[43], and the expressions $f_{xx}(\omega)$ and $f_{yy}(\omega)$ are the individual spectra of time series X_t and Y_t.

The squared coherence function can be interpreted as the frequency equivalent to the conventional squared correlation or the R-squared measure in the time domain. This motivates the interpretation of the coherency as a frequency measure of the degree to which series Y_t can be represented as a linear function of series X_t.

Another important measure in cross-spectral analysis is the phase spectrum, which measures the phase difference between the frequency components of processes X_t and Y_t. It is the frequency equivalent to the time lag in the time domain and is defined as:

$$\phi_{xy}(\omega) = \arctan\left(-\frac{q_{xy}(\omega)}{c_{xy}(\omega)} \right) \qquad (4.25)$$

The so defined phase spectrum indicates leading behaviour from process Y_t (leading indicator) to X_t (production) if $\phi_{xy}(\omega) < 0$, and respectively there is a lagged relation occurring if $\phi_{xy}(\omega) > 0$. When one variable is leading the other, $\phi_{xy}(\omega)/\omega$ measure the extent of the time lag (Granger 1969). The plot of the phase spectrum ϕ_{xy} over all feasible frequencies ω is called the phase diagram.

---

[43] Note that $f_{yx}(\omega) = \overline{f_{xy}(\omega)}$ holds, because of the relation of the covariance functions,

i.e. $\gamma_{yx}(h) = \gamma_{xy}(-h)$.

Coherence and phase are the tools, which will be used in the following empirical analysis in order to determine the relation of production and leading indicators. Together with the auto-spectrum of each series these will give an insightful view on the frequency interaction in the given framework.

Now that we have finished the short introduction to spectral estimation and cross-spectral methods, we can advance to apply the theory to our framework.

4.4. Spectral estimation and results

Before starting the analysis in the spectral domain, we have to take a look at the data series, which are used, because it has to be determined if they fulfil the required properties, i.e. if the series are demeaned and detrended. If not, there are some difficulties in applying spectral estimation. The problem of using a series with non-zero sample mean is that there is an abrupt offset when the time series is padded with zeros for the fast Fourier transform[44]. On the other side any trend component in the series would produce a spectral peak at zero frequency, which may dominate the estimation such that other important features are getting lost. In our case we already solved these problems in chapter 2 by taking first differences in order to ensure stationarity. Because of the transformation all time series have zero mean and possible trend components have been removed. Therefore we can start the spectral analysis by estimating the Fourier transform (4.9) and the raw periodogram (4.11). As already mentioned before, the raw periodogram is only a rough estimate of the spectrum and is not very useful for interpretation, because of its high variance of the spectral estimates. Therefore I will use the smoothing method introduced in section 4.2.4. in order to get a more stable spectral estimate.

---

[44] The fast Fourier transform (FFT) is the computational algorithm to estimate the discrete Fourier transform (DFT), whereby zeros are added before computing the spectral estimators.

Bloomfield (2000) recommends the so called Daniell filter to smooth the spectrum, whereby the resulting Daniell window of length m, is given by

$$g_i = \frac{1}{2(m-1)} \qquad \text{for } i = 1 \text{ and } i = m$$

and

$$g_i = \frac{1}{m-1} \qquad \text{otherwise.}$$

with m being the number of weights, while g_i represents the ith weight of the Daniell filter. Note that the first and the last weight are only half as large as the others, which leads to the typical trapezoid form of the weights. Using increasing lengths of Daniell filters is a popular method to estimate an increasingly smooth periodogram. One advantage of the Daniell filter is the relatively low leakage, which occurs because of the influence of the variance at non-Fourier frequencies on the spectrum at the Fourier frequencies.

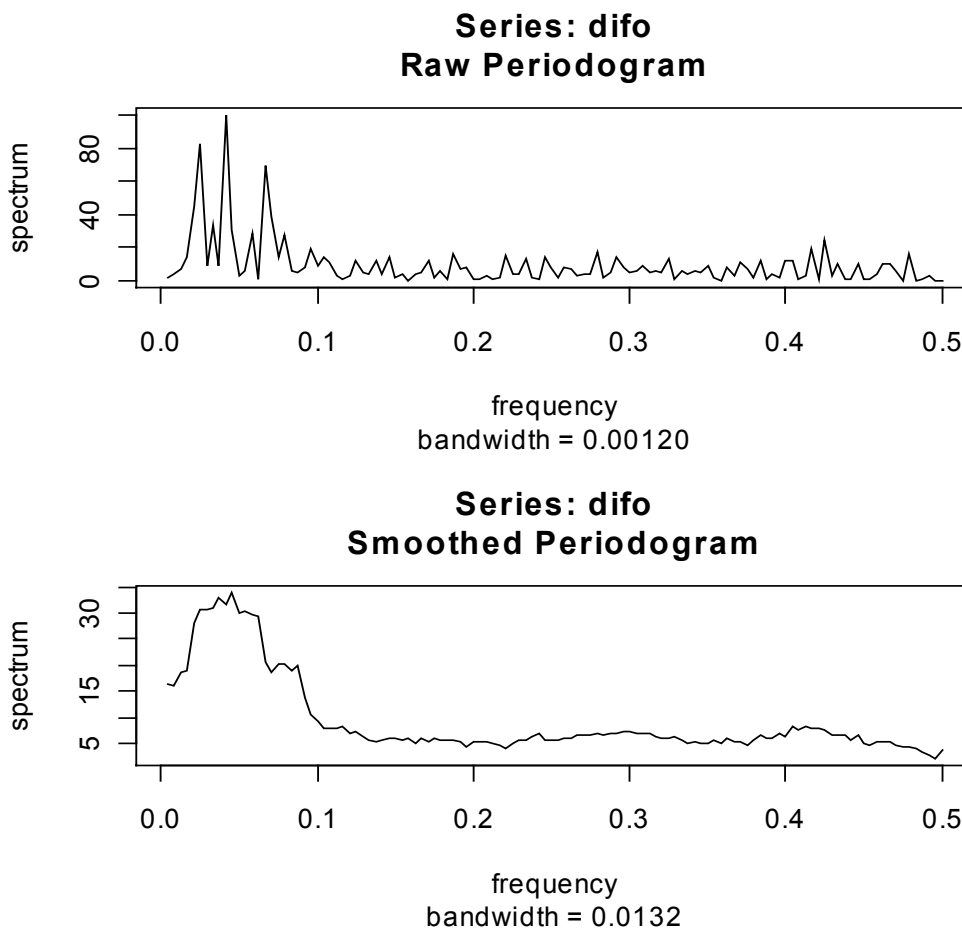**Figure 4.1.** *Raw and Smoothed periodogram of Ifo Business expectations (first differences → difo).*

Figure 4.1. shows the raw and the smoothed periodogram for the Ifo Business expectations series. The plots should give a visual example for the better understanding of the estimation procedure. In the given example it seems that the raw periodogram is a very noisy estimate of the theoretical spectrum. Its interpretation is a lot harder because of the high fluctuations. The smoothed periodogram seems to be a convenient compromise between stability and accuracy. The difficulty of finding the right frequency band is to ensure a stable estimation without losing important information. For the example given in Figure 4.1. I used the already explained Daniell filter and tried different values of m, starting with a low value and increasing the number until the desired stability was obtained. I have finally chosen a value of $m = 5$, which seemed to be reasonable, because at this point the spectrum looked quite stable without losing too much of the original shape of the raw periodogram and higher values, i.e. $m > 5$, did not change the estimation by much.

I performed the same procedure for all other series, which are used for our analysis[45]. These smoothed versions are taken for the final cross-spectral analysis. The estimation of coherence and phase should give a more sophisticated view on the relation of leading economic indicators and production.

Figure 4.2. *Squared coherency and phase spectrum of German production (logarithmic differences → dlprod_ger) and Ifo Business expectations (first differences → difo).*
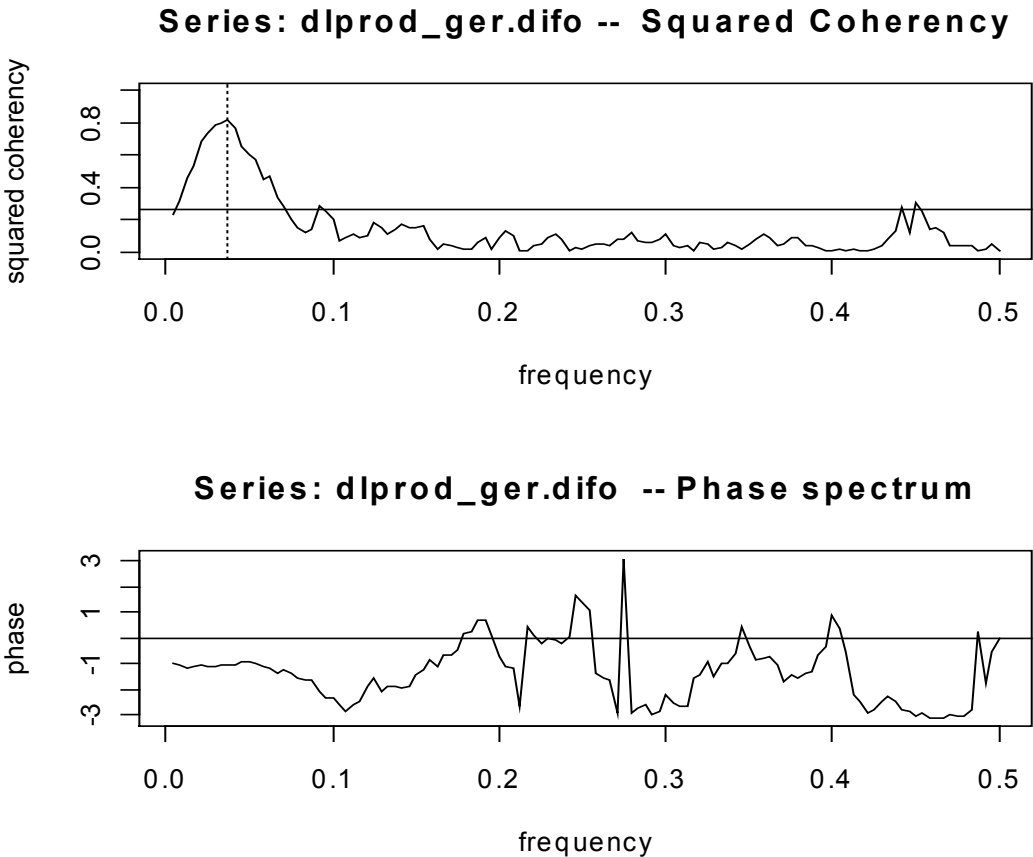
In section 4.3. we defined the squared coherency as a measure of the relation between two signals at different frequencies, similar to the R-squared measure in the time domain. Figure 4.2. shows the squared coherency for German production and Ifo Business expectations, whereby high values suggest a strong relation at the given frequency. The baseline in the figure represents the critical value of a simple F-test for significance at the 5% level, such that the null hypothesis of zero coherence can be rejected at a squared coherence value above the critical value.

---

[45] i.e. production account series, DAX and Consumer Confidence series.

The interpretation of the squared coherency plot between German production and Ifo Business expectations is quite interesting. There seems to be a strong relation at low frequencies (corresponding to long-run components). The highest estimate for the squared coherency reaches more than 80% at $\omega = 0.037$, which can be translated to a cycle every 27 months. After this peak the coherence decreases rapidly and finally becomes insignificant. Only a few frequencies around $\omega = 0.45$ are statistically significant. Therefore we can assume that (although there is strong long run relation) the predictive value of Ifo business expectations for German production is very limited at high frequency components (corresponding to the fast moving short run fluctuations). This result is in line with the findings of Lemmens/Croux/Dekimpe (2008), who targeted the performance of the European production expectation series, which are published by the Directorate General Economy and Finance of the European Union. These series are available for all EU member states and are constructed in a similar fashion to the German Ifo business expectations series. The authors analysed the causal relation of expectations and production in 12 European countries[46]. Lemmens/Croux/Dekimpe (2008) discovered that although most countries' production expectations are found to have highly significant (incremental) predictive power with respect to the longer-run components in the production account series, they have much more difficulty in predicting the fast-moving components of the production series[47]. Therefore it seems that surveys like the Ifo business expectation or the European production expectation series are not the best way to predict the upcoming economic situation, because of their deficits to explain the short run fluctuations in production, which should be the main task of leading economic indicators.

In the lower panel the phase spectrum gives some evidence for a leading behaviour of the Ifo business expectations on German production. This can be observed, because the values in the phase diagram are negative for nearly the entire frequency band[48]. The largest phase differences seem to be in the low and very high frequency components. At frequencies between 0.2 and 0.3 (corresponding to cycles between three and five months) the phase is shifting between positive and negative values, which makes the interpretation complicated. The most important result is that there seems to be leading behaviour of the Ifo business expectation series at the lower frequency components, where a strong relation between

---

[46] Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, The Netherlands and United Kingdom.

[47] See Lemmens, Croux, Dekimpe – Measuring and testing Granger causality over the spectrum: An application to European production expectation surveys (2008).

[48] See part 4.3. for definition and interpretation of the phase spectrum.

German production and Ifo business expectations was proved by estimating the squared coherency.

It should also be mentioned that our findings in the frequency domain are not inconsistent with the findings in the time domain. Remember that the time domain analysis suggested the Ifo business expectations to be a good and valid predictor for production. The lack of explaining the short run, which has been investigated in the frequency analysis, does not necessarily imply non-causality in an overall (time domain) Granger causality test. The Granger Wald test in the previous chapter indicates a strong causal relation, because of the highly significant coherence in the lower frequency components. This effect is stronger than the weak coherence in the fast oscillating parts. Therefore tests in the time domain, which do not distinguish between different frequency components, can have a different result than their frequency domain counterparts. This shows the importance of spectral decomposition. Without the use of spectral methods, we would not be able to identify the deficit of the Ifo business expectations in predicting the short run variations in German production.

The results for the other analysed economic leading indicators (i.e. consumer confidence and DAX) are similar. Although there is a rather strong coherence in the low frequency components, both indicators are mostly failing to explain the short and medium run oscillation in German production. The squared coherency plots are presented in Figure 4.3. and 4.4.

The squared coherency for consumer confidence and German production has its peak at $\omega = 1/48$, which is equivalent to a cycle every four years. At this frequency the squared coherency reaches the maximum value of more than 60%. Note that this predominant cycle is at a much lower frequency than for Ifo Business expectation series, where the maximum of 80% was reached at one cycle every 27 months. Another difference is the existence of significant frequency bands in the medium run, one around $\omega = 0.15$, another at $\omega = 0.21$. Beside these significant relations in the medium run, there seems to be no short run coherency.

The phase spectrum has a very interesting shape. At higher frequencies the production series seems to lead consumer confidence. This seems to be consistent with the results in the time domain, where we noticed a causal effect from German production to consumer confidence. The phase spectrum gives more evidence for this feedback relation. Therefore the use of consumer confidence as a leading economic indicator should be regarded as problematic.

Figure 4.3. *Squared coherency and phase spectrum of German production (logarithmic differences → dlprod_ger) and Consumer Confidence (first differences → dcons).*

Series: dlprod_ger.dcons -- Squared Coherency
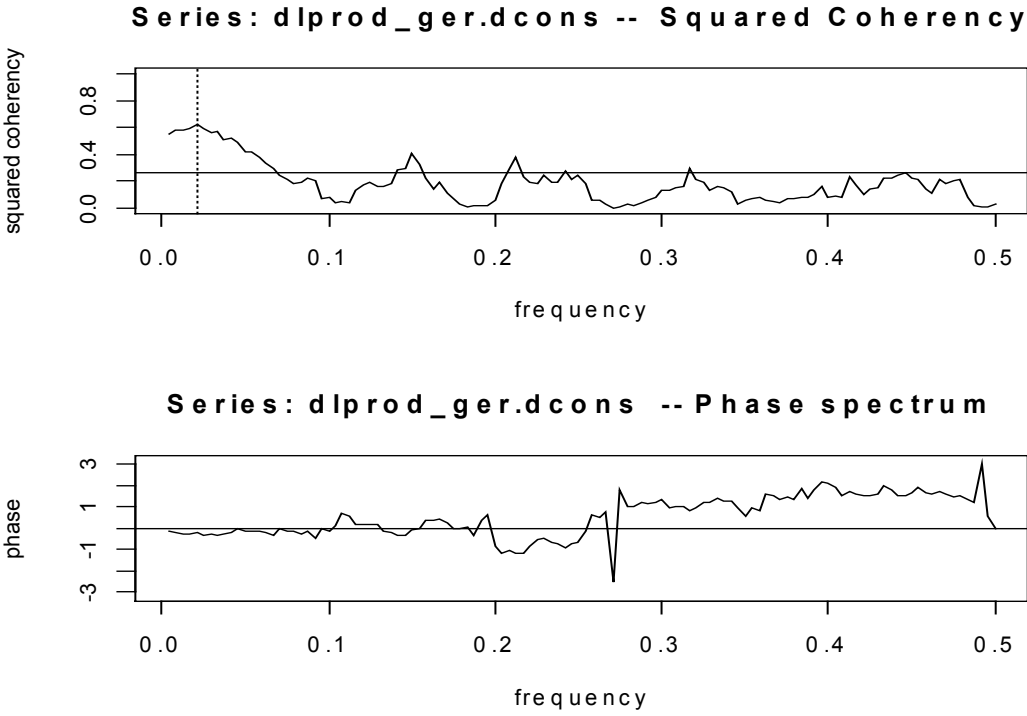

Series: dlprod_ger.dcons -- Phase spectrum

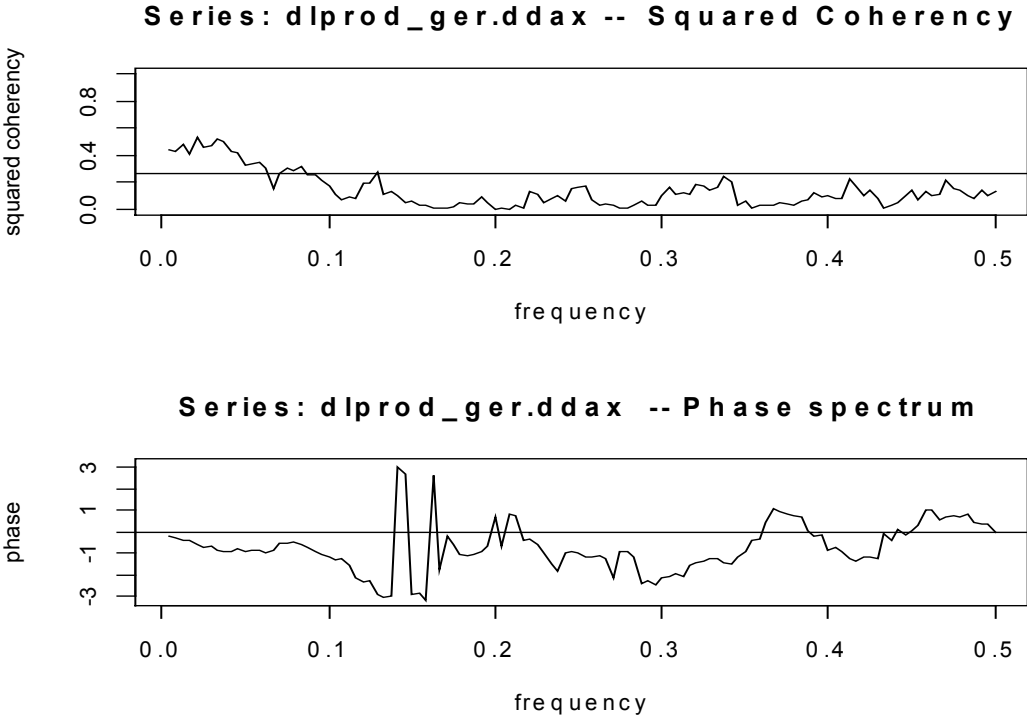***Figure 4.4.*** *Squared coherency and phase spectrum of German production (logarithmic differences → dlprod_ger) and DAX (first differences → ddax).*

Series: dlprod_ger.ddax -- Squared Coherency

Series: dlprod_ger.ddax -- Phase spectrum

49

Finally Figure 4.4. gives the squared coherency for German share index (DAX) and German production. Like for the other indicators there exists a low frequency coherency, but in a weaker form. The maximum coherency reaches only 50% at $\omega = 0.033$, and the relation declines more rapidly as the frequency gets higher. Only a rather small low frequency band is statistically significant and there is no real evidence for medium or short run coherence. The phase spectrum suggests leading behaviour of the German stock index for almost the entire long and medium run components, although there are some phase shifts in the frequency band between 0.15 and 0.2 (corresponding to cycles between five and six months). The interpretation at high frequencies is tricky, because of the up and down movement starting at $\omega = 0.35$. In this band there also exists evidence for leading behaviour of production on German stock index.

Like before[49], I have also analysed the spectral performance of the Ifo business expectations as a leading indicator for Austrian and European production. The time domain analysis suggested that the Ifo Business expectation series is a valid predictor for both production accounts. The frequency domain analysis shows a similar picture as for German production. The phase diagrams show the leading character of Ifo business expectations on Austrian and European production in the long run. The evidence for leading behaviour decreases for higher frequencies, because of the larger phase shifts in these components.

The coherence analysis between Ifo Business expectations and European/Austrian production confirms the same deficits as before in the German case. Although there is a statistically significant relation at the low frequency band, Ifo Business expectations cannot explain short run variations in Austrian or European production accounts. This result is straightforward, because there seems to be no reason, why the performance of Ifo Business expectations should be better for the foreign economic development.

---

[49] i.e. in the time domain analysis.

4.6. Conclusions for chapter 4

- There exists a low frequency relation between all analysed leading economic indicators and production. The strongest relation can be found between Ifo business expectations and German production.

- Although all leading indicators are viable in the low frequency parts (corresponding to long run movement), they have deficits in explaining the fast moving short run fluctuation in production account series.

- These shortcomings are problematic, because in theory the leading economic indicators are mainly developed and used for short run prediction.

- The findings in the frequency domain prove that positive results in the time domain causality tests are driven by low frequency coherence. Therefore the spectral analysis shows the limited predictive power of economic indicators over the spectrum, which is an insight that would not be recognised when using only time domain methods.

- The investigation of the phase diagrams gives some evidence for leading behaviour of the Ifo Business expectations and German stock index on production. Ifo Business expectations seem to lead production nearly over the entire frequency band, while DAX seems to lead production at low and medium frequency components.

- There is no evidence for leading behaviour of Consumer Confidence on German production. Moreover the phase diagram indicates that Consumer Confidence is lagging behind at high frequencies.

5. Summary and Conclusions

The purpose of my diploma thesis was the analysis of leading economic indicators, which have become major instruments for predicting future developments within the economy. I have discussed the characteristics of leading indicators and distinguished between specially constructed survey based indicators (Ifo Business expectations, Consumer confidence) and "natural" indicators (German stock index).

Different mathematical methods have been used to investigate the performance of all these measures, whereby the focus was mostly on the Ifo Business expectation survey, which seems to be the most important leading indicator in Europe.

In chapter three I used bivariate VAR-models to analyse the causal relation between leading indicators and production in the time domain. Therefore I introduced the concepts of Granger Causality (GC), feedback and instantaneous causality. Testing procedures for all these measures have been discussed in detail. The test results gave a first impression of the predictive power of leading indicators in the time domain. It could be proved that all measures are Granger-causing German production, whereas Ifo Business expectations displayed the best statistical properties. The results for the Consumer confidence survey caused some problems due the existence of feedback. Therefore it was not possible to determine the causal direction between Consumer confidence and German production, which led to the conclusion that demand side indicators might not be viable leading indicators.

I also dealt with the question whether Ifo Business expectations are able to predict Austrian and European[50] production. In both cases I was able to establish a causal relation, which proves Germany's huge impact on European economy. Nevertheless, I suggested that these results ought to be handled with care because of the feedback mechanism that was identified between Ifo Business expectations and European/Austrian production.

After the analysis in the time domain, I gave a short theoretical introduction to spectral estimation of economic time series. I explained the decomposition of stationary series into their basic frequency components and the estimation of cross-spectra. Afterwards I used the concepts of squared coherency and phase diagrams to analyse the relation of economic leading indicators and production in the frequency domain. The use of spectral methods exhibited various implications. On one side these methods can give a more insightful view by formally decomposing the data in their basic components, on the other side I wanted to show

---

[50] i.e. the production in the euro zone.

an alternative way to the well-known time domain analysis. Of course the usefulness of time series models is undeniable for empirical research, but one should also take into account that there are other concepts, which can be used for further analysis and brighten the view on the research topic.

In our case, the analysis in the frequency domain showed some interesting features of the relation between leading economic indicators and production. Despite a significant relation[51] in the low frequency components (corresponding to long run movement), all tested leading indicators had deficits in explaining higher frequency oscillation in production accounts. These results in the frequency domain indicate that the causality in the time domain is mostly driven by the low frequency coherence.

The investigation of the phase diagrams showed some evidence for leading behaviour of Ifo Business expectations and German stock index on production, while the Consumer Confidence could not be identified as a leading indicator, which was in line with the time domain analysis.

The main lesson that should be drawn from this diploma thesis is that although leading economic indicators seem to have good predictive power when using testing procedures in the time domain, limited predictive power can be noticed when applying spectral methods. The shortcomings of all analysed measures[52] in explaining short run variations are a problem that should not be ignored because leading indicators like the Ifo Business expectations and the Consumer confidence are mainly constructed to predict short run development. If they are not able to give additional information about short run behaviour in production accounts, one could doubt the necessity of spending so much money on the conduct of large surveys. As a matter of fact it seems that Niels Bohr was right. Predictions are difficult, especially about the future.

---

[51] i.e. in terms of squared coherency.
[52] i.e. Ifo Business expectations, Consumer confidence and DAX.

6. Literature

Akaike H. (1969), Fitting autoregressive models for prediction, Annals of the Institute of Statistical Mathematics 23, 163-180.

Ashley R., C. W. J. Granger and R. Schmalensee (1980), Advertising and Aggregate Consumption: An Analysis of Causality, Econometrica, 48, 1149-1167.

Bierbaum-Polly J. (2010), Composite Leading Indicator for the Austrian Economy: Methodology and "Real-time" Performance, WIFO Working Papers, No. 369.

Bloomfield P. (2000), Fourier Analysis of time series: An introduction (Second Edition), Wiley, New York.

Breitung J. & Candelon B. (2006), Testing for short- and long-run causality: A frequency domain approach. Journal of Econometrics, 132, 363-378.

Brockwell P.J., and Davis R.A. (1996), Introduction to Time Series and Forecasting, Springer-Verlag New York.

Brockwell P.J., and Davis R.A. (1991), Time Series: Theory and Methods, Springer-Verlag New York.

Burns A.F. and W.C. Mitchell (1946), Measuring Business Cycles, National Bureau of Economic Research.

Dickey, D.A. and W.A. Fuller (1979), Distribution of the Estimators for Autoregressive Time Series with a Unit Root, Journal of the American Statistical Association, 74, 427-431.

Geweke J. (1986), Measurement of linear dependence and feedback between multiple time series, Journal of the American Statistical Association, 77, 304-324

Granger C. W. J. and M. Hatanka (1964), Spectral Analysis of Economic Time Series, Princeton University Press, Princeton, New Jersey.

Granger C. W. J. (1966), The Typical Spectral Shape of an Economic Variable, Econometrica, Vol. 34(1), pp.150-161.

Granger C. W. J. (1980), Testing for Causality, Journal of Economic Dynamics and Control, 2, 329-352.

Granger C. W. J. (1969), Investigating Causal Relations by Econometric Models and Cross-Spectral Methods, Econometrica, Vol. 37(3), pp. 424-438.

Greene W. H. (2000), Econometric Analysis (Fourth Edition), Prentice-Hall.

Iacobucci A. (2003), Spectral Analysis for Economic Time Series, OFCE.

Koopmans L. H. (1995), The spectral analysis of time series, Probability and Mathematical Statistics, Vol. 22, San Diego: Academic Press.

Lemmens A., C. Croux and M. G. Dekimpe (2008), Measuring and testing Granger causality over the spectrum: An application to European production expectation surveys, International Journal of Forecasting 24.

Lütkepohl H. (2004), Applied Time Series Econometrics, Cambridge University Press.

Lütkepohl H. (2006), New Introduction to Multiple Time Series Analysis, Springer-Verlag Berlin.

Pierce D.A. (1979), R-squared measures for time series, Journal of the American Statistical Association, 74, 901-910

Pierce D.A. (1982), The measurement of linear-dependence and feedback between multiple time series – Comment, Journal of the American Statistical Association, 77, 315.

Sees A. (2009), German leading indicators' ambitious message, Economics & FI/FX Research, Economic Special, UniCredit Group.

Shumway R. and D. Stoffer (2006), Time Series Analysis and Its Applications (Second Edition), Springer Science & Business Media.

Toda H.Y. and P.C.B. Phillips (1993), Vector autoregressions and causality, Econometrica 61, 1367-1393.

Toda H.Y. and T. Yamamoto (1995), Statistical inference in vector autoregressions with possibly integrated processes, Journal of Econometrics, 66, 225-250.

Appendix A: Abstract

A.1. English version

In recent years leading economic indicators have become more and more important. Nowadays, the entire media landscape follows the development of these measures and their information content is often used by economists whenever future predictions are needed. Despite the fact that leading economic indicators are heavily utilized, their real predictive power is almost uninvestigated. This provides the motivation for analysing whether leading economic indicators are really useful tools for forecasting the upcoming economic development.

The present diploma thesis covers and analyses some of the most important leading indicators in Europe (Ifo Index, Consumer Confidence and German Stock Index). Various statistical methods are used in order to test if there is a link between these measures and the upcoming production within the economy.

In the first part of the analysis the well-known time series models are implemented. In the process the concept of Granger Causality is applied for investigating the existence of a causal relation between leading economic indicators and production in the time domain.

Furthermore, in the second part spectral methods are utilized to get a more insightful view on the topic. These methods allow the formal decomposition of the data in their basic components, which enables me to investigate the relation of leading indicators and production over different frequencies. In this way it is possible to determine the exact frequency bands in which leading indicators contain useful information for the explanation of future economic developments.

A.2. German version

In den letzten Jahren gewannen konjunkturelle Vorlaufindikatoren immer stärker an Gewicht. Sowohl in medialer als auch wissenschaftlicher Hinsicht sind sie zu wichtigen Werkzeugen für die Erstellung von Wirtschaftsprognosen geworden. Die Frage ob und wie gut konjunkturelle Vorlaufindikatoren wirklich zum besseren Verständnis zukünftiger Ereignisse beitragen können, ist allerdings weitgehend unerforscht. Dieser Umstand liefert die Motivation zur Analyse der Prognosefähigkeit von ausgewählten Indikatoren um dadurch die Sinnhaftigkeit ihrer Erstellung genauer zu hinterfragen.

In der vorliegenden Diplomarbeit werden einige der wichtigsten Vorlaufindikatoren im europäischen Raum (Ifo Index, Konsumentenvertrauen und DAX) erläutert und analysiert. Mittels verschiedener statistischer Modelle wird getestet ob ein kausaler Zusammenhang zwischen den jeweiligen Vorlaufindikatoren und der Produktionsleistung innerhalb der Volkswirtschaft besteht.

Um die Kausalitätsfrage zu beantworten werden zunächst gängige Zeitreihenmodelle (time domain) verwendet. Dabei wird das Konzept der sogenannten Granger-Kausalität auf die Fragestellung der Prognosefähigkeit von konjunkturellen Vorlaufindikatoren angewandt.

Im zweiten Teil der Arbeit werden zusätzlich Methoden der Spektralanalyse herangezogen, Diese ermöglichen es den Zusammenhang zwischen konjunkturellen Vorlaufindikatoren und Produktion über verschiedene Frequenzintervalle zu bestimmen (frequency domain). Dadurch soll genauer untersucht werden, in welchen Frequenzkomponenten der Vorlaufindikatoren nützliche Informationen für die Erklärung zukünftiger wirtschaftlicher Entwicklungen enthalten sind.

Appendix B: Curriculum Vitae

Name:

Michael Wüger

Date of birth:

October 27th, 1984

Place of birth:

Vienna, Austria

Education:

2005-2010:   Studies of Economics at the University of Vienna, Austria

2009:   ERASMUS-studies at the Free University of Berlin (FU Berlin), Germany

1995-2003:   Gymnasium Neusiedl/See, Austria

Language Skills:

German: Mother tongue

English: First foreign language, fluent in spoken and written

French: Second foreign language, six years of school education